



Methodological challenges and solutions in auditory functional magnetic resonance imaging

Jonathan E. Peelle *

Department of Otolaryngology, Washington University in St. Louis, St. Louis, MO, USA

Edited by:

Pedro Antonio Valdes-Sosa, Centro de Neurociencias de Cuba, Cuba

Reviewed by:

Sibylle C. Herholz, Deutsches Zentrum für Neurodegenerative Erkrankungen, Germany
Dezhong Yao, University of Electronic Science and Technology of China, China

*Correspondence:

Jonathan E. Peelle, Department of Otolaryngology, Washington University in St. Louis, 660 South Euclid, Box 8115, St. Louis, MO 63110, USA
e-mail: peellej@ent.wustl.edu

Functional magnetic resonance imaging (fMRI) studies involve substantial acoustic noise. This review covers the difficulties posed by such noise for auditory neuroscience, as well as a number of possible solutions that have emerged. Acoustic noise can affect the processing of auditory stimuli by making them inaudible or unintelligible, and can result in reduced sensitivity to auditory activation in auditory cortex. Equally importantly, acoustic noise may also lead to increased listening effort, meaning that even when auditory stimuli are perceived, neural processing may differ from when the same stimuli are presented in quiet. These and other challenges have motivated a number of approaches for collecting auditory fMRI data. Although using a continuous echoplanar imaging (EPI) sequence provides high quality imaging data, these data may also be contaminated by background acoustic noise. Traditional sparse imaging has the advantage of avoiding acoustic noise during stimulus presentation, but at a cost of reduced temporal resolution. Recently, three classes of techniques have been developed to circumvent these limitations. The first is Interleaved Silent Steady State (ISSS) imaging, a variation of sparse imaging that involves collecting multiple volumes following a silent period while maintaining steady-state longitudinal magnetization. The second involves active noise control to limit the impact of acoustic scanner noise. Finally, novel MRI sequences that reduce the amount of acoustic noise produced during fMRI make the use of continuous scanning a more practical option. Together these advances provide unprecedented opportunities for researchers to collect high-quality data of hemodynamic responses to auditory stimuli using fMRI.

Keywords: auditory cortex, auditory perception, speech, music, hearing, executive function

INTRODUCTION

Over the past 20 years, functional magnetic resonance imaging (fMRI) has become the workhorse of cognitive scientists interested in noninvasively measuring localized human brain activity. Although the benefits provided by fMRI have been substantial, there are numerous ways in which it remains an imperfect technique. This is perhaps nowhere more true than in the field of auditory neuroscience due to the substantial acoustic noise generated by standard fMRI sequences. In order to study brain function using fMRI, auditory researchers face what can seem like an unappealing array of methodological decisions that impact the acoustic soundscape, cognitive performance, and imaging data characteristics to varying degrees. Here I review the challenges faced in auditory fMRI studies, possible solutions, and prospects for future improvement. Much of the information regarding the basic mechanics of noise in fMRI can be found in previous reviews (Amaro et al., 2002; Moelker and Pattynama, 2003; Talavage et al., 2014); although I have repeated the main points for completeness, I focus on more recent theoretical perspectives and methodological advances.

SOURCES OF ACOUSTIC INTERFERENCE IN fMRI

Table 1 summarizes several factors that contribute to the degradation of acoustic signals during fMRI. Echoplanar imaging (EPI) sequences commonly used to detect the blood oxygen level dependent (BOLD) signal in fMRI require radiofrequency (RF) pulses that excite tissue and gradient coils that help encode spatial position by altering the local magnetic field. During EPI the gradient coils switch between phase encoding and readout currents, producing Lorentz forces that act on the coils and connecting wires. These vibrations travel as compressional waves through the scanner hardware and eventually enter the air as acoustic sound. This gradient-induced vibration produces the most prominent acoustic noise during fMRI, and can continue for up to approximately 0.5 s after the gradient activity ceases (Ravicz et al., 2000). Because the Lorentz force is proportional to the main magnetic field strength (B_0) and the gradient current, both high B_0 and high gradient amplitudes generally increase the amount of acoustic noise generated (Moelker et al., 2003). For example, increasing field strength from 0.2 to 3 T will bring maximum acoustic noise from ~85 to ~130 dB SPL (Foster et al., 2000; Ravicz et al., 2000; Price et al., 2001).

Table 1 | Sources of acoustic interference during fMRI.

| Source | Approximate noise level (dB SPL) |
|---------------------------------|----------------------------------|
| Gradient coils | 85–130 |
| Helium pump and air circulating | 57–76 |
| In-ear foam earplugs | – |
| Sub-optimal headphones | – |

Although the noise generated by gradient switching is the most obvious (i.e., loudest) source of acoustic noise during fMRI, it is not the only source of acoustic interference. RF pulses contribute additional acoustic noise, and noise is also present as a result of air circulation systems and helium pumps in the range of 57–76 dB SPL (Ravicz et al., 2000). Because RF and helium pump noise is substantially quieter than that generated by gradient coils it probably provides a negligible contribution when scanning is continuous, but may be more relevant in sparse or interleaved silent steady state (ISSS) imaging sequences (described in a later section) when gradient-switching noise is absent. Auditory clarity can also be reduced as a result of in-ear hearing protection and sub-optimal headphone systems.

Separately or together, these noise sources provide a level of acoustic interference that is significantly higher than that found in a typical behavioral testing environment. In the next section I turn to the more interesting question of the various ways in which this cacophony may impact auditory neuroscience.

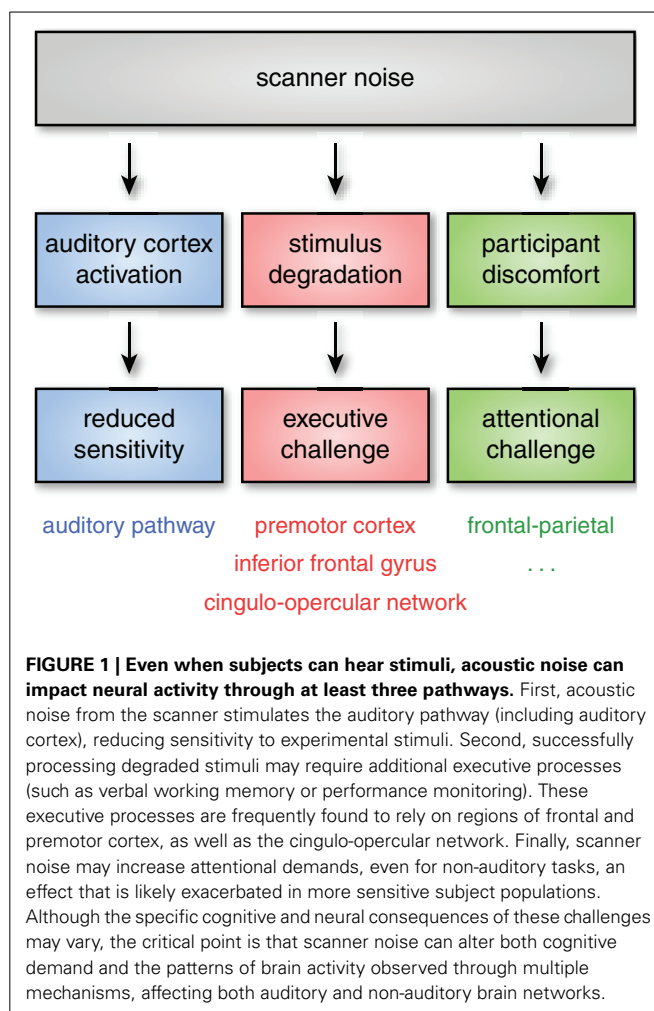
CHALLENGES OF ACOUSTIC NOISE IN AUDITORY fMRI

Acoustic noise can influence neural response through at least three independent pathways, illustrated schematically in **Figure 1**. The effects will vary depending on the specific stimuli, population being studied, and brain networks being examined. Importantly, though, in many cases the impact of noise on brain activation can be seen outside of auditory cortex. In this section I review the most pertinent challenges caused by acoustic scanner noise.

ENERGETIC MASKING

Energetic masking refers to the masking of a target sound by a noise or distractor sound that obscures information in the target. That is, interference occurs at a peripheral level of processing, with the masker already obscuring the target as the sound enters the eardrum (and thus at the most peripheral levels of the auditory system). The level of masking is often characterized by the signal-to-noise ratio (SNR), which reflects the relative loudness of the signal and masker. For example, an SNR of +5 dB indicates that on average the target signal is 5 dB louder than the masker. If scanner noise at a subject's ear is 80 dB SPL, achieving a moderately clear SNR of +5 would require presenting a target signal at 85 dB SPL. When considering the masking effects of noise it is important to note that the characteristics of the noise are also important: noise that has temporal modulation can permit listeners to glean information from the “dips” in the noise masker.

Energetic masking highlights the most obvious challenge of using auditory stimuli in fMRI: Subjects may not be able to perceive auditory stimuli due to scanner noise. If stimuli are inaudible—or less than fully perceived in some way—interpreting the subsequent neural responses can be problematic. A different



(but related) sort of energetic masking challenge arises in experiments in which subjects are required to make vocal responses, as scanner noise can interfere with an experimenter's understanding of subject responses; in some cases this can be ameliorated by offline noise reduction approaches (e.g., Cusack et al., 2005). In addition, the presence of acoustic noise may also change the quality of vocalizations produced by subjects (Junqua, 1996). Acoustic noise thus impacts not only auditory perception, but speech production, which may be important for some experimental paradigms.

Two ways of ascertaining the degree to which energetic masking is a problem are (1) to ask participants about their subjective experience hearing stimuli or (2) to include a discrimination or recall test that can empirically verify the degree to which auditory stimuli are perceived. Given individual differences in hearing level and ability to comprehend stimuli in noise, these are likely best done for each subject, rather than, for example, audibility being verified solely by the experimenter. It is also important to test audibility using stimuli representative of those used in the experiment, as the masking effects of scanner noise can be influenced by specific acoustic characteristics of the target stimuli (for example, being more detrimental to perception of birdsong than speech).

Although it is naturally important for subjects to be able to hear experimental stimuli (and for experimenters to hear subject responses, if necessary), the requirement of audibility is obvious enough that it is often taken into account when designing a study. However, acoustic noise may also cause more pernicious challenges, to which I turn in the following sections.

AUDITORY ACTIVATION

A natural concern regarding acoustic noise during fMRI relates to the activation along the auditory pathway resulting from the scanner noise. If brain activity is modulated in response to scanner noise, might this reduce our ability to detect signals of interest? To investigate the effect of scanner noise on auditory activation, Bandettini et al. (1998) acquired data with and without EPI-based acoustic stimulation, enabling them to compare brain activity that could be attributed to scanner noise. They found that scanner noise results in increased activity bilaterally in superior temporal cortex (see also Talavage et al., 1999). Notably, this activity was not observed only in primary auditory cortex, but in secondary auditory regions as well. The timecourse of activation to scanner noise peaks 4–5 s after stimulus onset, returning to baseline by 9–12 s (Hall et al., 2000), and is thus comparable to that observed in other regions of cortex (Aguirre et al., 1998). Scanner-related activation in primary and secondary auditory cortex limits the dynamic range of these regions, producing weaker responses to auditory stimuli (Shah et al., 1999; Talavage and Edmister, 2004; Langers et al., 2005; Gaab et al., 2007). In addition to overall changes in magnitude or spatial extent of auditory activation, scanner noise can affect the level at which stimuli need to be presented for audibility, which can in turn affect activity down to the level of tonotopic organization (Langers and van Dijk, 2012). Thus, if activity along the auditory pathway proper is of interest, the contribution of scanner noise must be carefully considered when interpreting results.

It is worth noting that while previous studies have investigated the effect of scanner noise on overall (univariate) response magnitude, the degree to which this overall change in gain may affect multivariate analyses is unclear. Again, this is true for activity in both auditory cortex and regions further along the auditory processing hierarchy (Davis and Johnsrude, 2007; Peelle et al., 2010b).

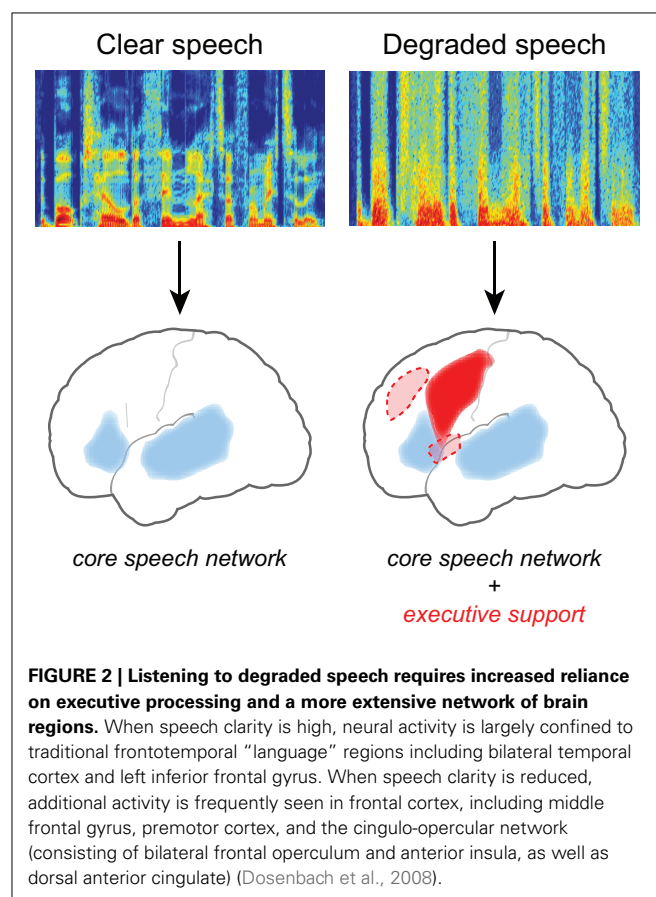
COGNITIVE EFFORT DURING AUDITORY PROCESSING

Although acoustic noise can potentially affect all auditory processing, most of the research on the cognitive effects of acoustic challenge has occurred in the context of speech comprehension. There is increasing consensus that decreased acoustic clarity requires listeners to engage additional cognitive processing to successfully understand spoken language. For example, after hearing a list of spoken words, memory is worse for words presented in noise, even though the words themselves are intelligible (Rabbitt, 1968). When some words are presented in noise (but are still intelligible), subjects have difficulty remembering not only the words in noise, but prior words (Rabbitt, 1968; Cousins et al., 2014), suggesting an increase in cognitive processing for degraded speech that lasts longer than the degraded stimulus itself and interferes with memory (Miller and Wingfield, 2010). Additional

evidence supporting the link between acoustic challenge and cognitive resources comes from pupillometry (Kuchinsky et al., 2013; Zekveld and Kramer, 2014) and visual tasks which relate to individual differences in speech perception ability (Zekveld et al., 2007; Besser et al., 2012). The additional cognitive resources required are not specific to acoustic processing but appear to reflect more domain-general processes (such as verbal working memory) recruited to help with auditory processing (Wingfield et al., 2005; Rönnberg et al., 2013). Thus, acoustic challenge can indirectly impact a wide range of cognitive operations.

Consistent with this shared resource view, behavioral effects of acoustic clarity are reliably found on a variety of tasks. Van Engen et al. (2012) compared listeners' recognition memory for sentences spoken in conversational speech compared to those spoken in a clear speaking style (with accentuated acoustic features), and found that memory was superior for the acoustically-clearer sentences. Likewise, listeners facing acoustic challenge—due to background noise, degraded speech, or hearing impairment—perform poorer than listeners with normal hearing on auditory tasks ranging from sentence processing to episodic memory tasks (Pichora-Fuller et al., 1995; Surprenant, 1999; Murphy et al., 2000; McCoy et al., 2005; Tun et al., 2010; Heinrich and Schneider, 2011; Lash et al., 2013).

Converging evidence for the neural effects of effortful listening comes from fMRI studies in which increased neural activity is seen for degraded speech relative to unprocessed speech (Scott and McGettigan, 2013), illustrated in **Figure 2**. Davis and



Johnsrude (2003) presented listeners with sentences that varied in their intelligibility, with speech clarity ranging from unintelligible to fully intelligible. They found greater activity for degraded speech compared to fully intelligible speech in the left hemisphere, along both left superior temporal gyrus and inferior frontal cortex. Importantly, increased activity in frontal and prefrontal cortex was greater for moderately distorted speech than either fully intelligible or fully unintelligible speech (i.e., an inverted U-shaped function), consistent with its involvement in recovering meaning from degraded speech (as distinct from a simple acoustic response). Acoustic clarity (i.e., SNR) also impacts the brain networks supporting semantic processing during sentence comprehension (Davis et al., 2011), possibly reflecting increased use of semantic context as top-down knowledge during degraded speech processing (Obleser et al., 2007; Obleser and Kotz, 2010; Sohoglu et al., 2012).

Additional studies using various forms of degraded speech have also found difficulty-related increases in regions often associated with cognitive control or performance monitoring, such as bilateral insula and anterior cingulate cortex (Eckert et al., 2009; Adank, 2012; Wild et al., 2012; Erb et al., 2013; Vaden et al., 2013). The stimuli used in these studies are typically less intelligible than unprocessed speech (e.g., 4- or 6-channel vocoded¹ speech, or low-pass filtered speech). Thus, although the increased recruitment of cognitive and neural resources to handle degraded speech is frequently observed, the specific cognitive processes engaged—and thus the pattern of neural activity—depend on the degree of acoustic challenge presented. An implication of this variability is that it may be hard to predict *a priori* the effect of acoustic challenge on the particular cognitive system(s) of interest.

In summary, there is clear evidence that listening to degraded speech results in increased cognitive demand and altered patterns of brain activity. The specific differences in neural activity depend on the degree of the acoustic challenge, and thus may differ between moderate levels of degradation (when comprehension accuracy remains high and few errors are made) and more severe levels of degradation (when comprehension is significantly decreased). It is important to note that effort-related differences in brain activity can be seen both within the classic speech comprehension network and in regions less typically associated with speech comprehension, and depend on the nature of both the stimuli and the task. Furthermore, the way in which these effort-related increases interact with other task manipulations has received little empirical attention, and thus the degree to which background noise may influence observed patterns of neural response for many specific tasks is largely unknown.

Finally, although most of the research on listening effort has been focused on speech comprehension, it is reasonable to think that many of these same principles might transfer to other auditory domains, such as music or environmental sounds. And,

as covered in the next section, effects of acoustic challenge need not even be limited to auditory tasks.

EFFECTS OF ACOUSTIC NOISE IN NON-AUDITORY TASKS

Although the interference caused by acoustic noise is most obvious when considering auditory tasks, it may also affect subjects' performance on non-auditory tasks (for example, by increasing demands on attention systems). The degree to which noise impacts non-auditory tasks is an important one for cognitive neuroscience. Unfortunately, there have been relatively few studies addressing this topic directly.

Using continuous EPI, Cho et al. (1998a) had subjects perform simple tasks in the visual (flickering checkerboard) and motor (finger tapping) domains, with and without additional scanner noise played through headphones. The authors found opposite effects in visual and motor modalities: activity in visual cortex was increased with added acoustic noise, whereas activity in motor cortex was reduced.

To investigate the effect of scanner noise on verbal working memory, Tomasi et al. (2005) had participants to perform an *n*-back task using visually-displayed letters. The loudness of the EPI scanning was varied by approximately 12 dB by selecting two readout bandwidths to minimize (or maximize) the acoustic noise. No difference in behavioral accuracy was observed as a function of noise level. However, although the overall spatial patterns of task-related activity were similar, brain activity differed as a function of noise. The louder sequence was associated with increased activity in several regions including large portions of (primarily dorsal) frontal cortex and cerebellum, and the quieter sequence was associated with greater activity in (primarily ventral) regions of frontal cortex and left temporal cortex.

Behaviorally, recorded scanner noise has been shown to impact cognitive control (Hommel et al., 2012); additional effects of scanner noise have been reported in fMRI tasks of emotional processing (Skouras et al., 2013) and visual mental imagery (Mazard et al., 2002). Thus, MRI acoustic noise influences brain function across a number of cognitive domains.

It is not only the loudness of scanner noise that is an issue, but also the characteristics of the sound: whether an acoustic stimulus is pulsed or continuous, for example, can significantly impact both auditory and attentional processes. Haller et al. (2005) had participants perform a visual *n*-back task, using either a conventional EPI sequence or one with a continuous sound (i.e., not pulsed). Although behavioral performance did not differ across sequence, there were numerous differences in the detected neural response. These included greater activity in cingulate and portions of frontal cortex for the conventional EPI sequence, but greater activity in other portions of frontal cortex and left middle temporal gyrus for the continuous noise sequence. As with conventional EPI sequences, scanner noise is once again found to impact neural processing in areas beyond auditory cortex (see also Haller et al., 2009).

It is worth noting that not every study investigating this issue has observed effects of acoustic noise in non-auditory tasks: Elliott et al. (1999), using participants performing visual, motor, and auditory tasks, found that scanner noise resulted in decreased activity uniquely during the auditory condition. Nevertheless, the

¹Noise vocoding (Shannon et al., 1995) involves dividing the frequency spectrum of a stimulus into bands, or channels. Within each channel, the amplitude envelope is extracted and used to modulate broadband noise. Thus, the number of channels determines the spectral detail present in a speech signal, with more channels resulting in a more detailed (and for speech, more intelligible) signal (see Figure 2 in Peelle and Davis, 2012).

number of instances in which scanner noise has been found to affect neural activity on non-auditory tasks is high enough that the issue should be taken seriously: Although exactly how much of the difference in neural response can be attributed to scanner noise is debatable, converging evidence indicates that the effects of scanner noise frequently extend beyond auditory cortex (and auditory tasks). These studies suggest that (1) a lack of behavioral effect of scanner noise does not guarantee equivalent neural processing; (2) both increases and decreases in neural activity are seen in response to scanner noise; and (3) the specific regions in which noise-related effects are observed vary across study.

OVERALL SUBJECT COMFORT AND SPECIAL POPULATIONS

An additional concern regarding scanner noise is that it may increase participant discomfort. Indeed, acoustic noise can cause anxiety in human subjects (Quirk et al., 1989; Meléndez and McCrank, 1993), a finding which may also extend to animals. Scanner noise presents more of a challenge for some subjects than others, and it may be possible to improve the comfort of research subjects (and hopefully their performance) by reducing the amount of noise during MRI scanning. Additionally, if populations of subjects differ in a cognitive ability such as auditory attention, the presence of scanner noise may affect one group more than another. For example, age can significantly impact the degree to which subjects are bothered by environmental noise (Van Gerven et al., 2009); similarly, individual differences in noise sensitivity may contribute to (or reflect) variability in the effects of scanner noise on neural response (Pripfl et al., 2006). These concerns may be particularly relevant in clinical or developmental studies with children, participants with anxiety or other psychiatric condition, or participants who are particularly bothered by auditory stimulation.

A CAUTIONARY NOTE REGARDING INTERACTIONS

One argument sometimes made in auditory fMRI studies using standard EPI sequences is that although acoustic noise may have some overall impact, because noise is present during all experimental conditions it cannot influence the results when comparing across conditions (which is often of most scientific interest). Given the ample amount of evidence for auditory-cognitive interactions, such an assumption seems tenuous at best. If anything,

there is good reason to suspect interactions between acoustic noise and task difficulty, which may manifest differently depending on particular stimuli, listeners, and statistical methods (for example, univariate vs. multivariate analyses). In the absence of empirical support to the contrary, claims that acoustic noise is unimportant should be treated with skepticism.

SOLUTIONS FOR AUDITORY fMRI

Although at this point the prospects for auditory neuroscience inside an MRI scanner may look bleak, there is still cause for optimism. In this section I provide an overview of several methods for dealing with scanner noise that have been employed, noting advantages and disadvantages of each. These approaches are listed in Table 2, a subset of which is shown in Figure 3.

PASSIVE HEARING PROTECTION

Subjects in MRI studies typically wear over-ear hearing protection that attenuates acoustic noise by approximately 30 dB. Subjects may also wear insert earphones, or foam earplugs that can provide additional reduction in acoustic noise of 25–28 dB, for a combined reduction of approximately 40 dB (Ravicz and Melcher, 2001). Although hearing protection can reduce the acoustic noise perceived during MRI, it cannot eliminate it completely: Even if perfect acoustic isolation could be achieved at the outer ear, sound waves still travel to the cochlea through bone conduction. Thus, hearing protection is only partial solution, and some degree of auditory stimulation during conventional fMRI is unavoidable. In addition, passive hearing protection may change the frequency spectrum of stimuli, affecting intelligibility or clarity.

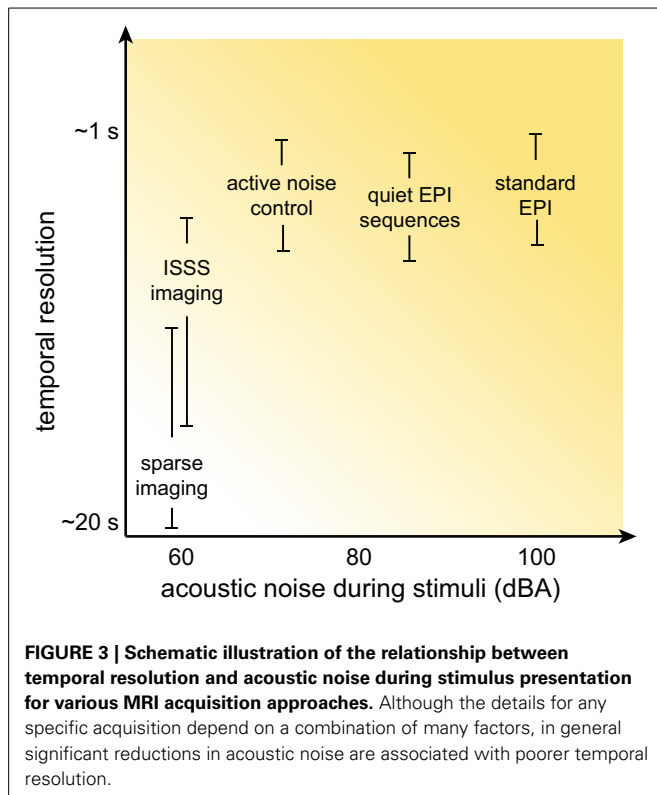
CONTINUOUS SCANNING USING A STANDARD EPI SEQUENCE

One approach in auditory fMRI is to present stimuli using a conventional continuous scanning paradigm, taking care to ensure that participants are able to adequately hear the stimuli (Figure 4A). This approach generally assumes that, because scanning noise is consistent across experimental condition, it is unlikely to systematically affect comparisons among conditions (typically what is of interest). I have already noted above the danger of this assumption with respect to additional task effects and ubiquitous interactions between perceptual and cognitive factors. However, for some paradigms a continuous scanning paradigm

Table 2 | Methods for dealing with acoustic noise in fMRI.

| Approach | Approximate noise reduction during stimulus (dB) ^a | Requires custom scanner hardware? | Requires custom presentation equipment? | Requires custom MRI sequence? | Image quality relative to continuous | Temporal resolution relative to continuous |
|-------------------------------|---|-----------------------------------|---|-------------------------------|--------------------------------------|--|
| Continuous EPI | 0 | No | No | No | – | – |
| Passive hearing protection | 35 | No | No | No | No change | No change |
| Sparse imaging | 50 | No | No | No | No change | Reduced |
| ISSS imaging | 50 | No | No | Yes | No change | Slightly reduced |
| Active noise control | 40 | No | Yes | No | No change | No change |
| Quiet MRI sequences | 20 | No | No | Yes | Reduced | Slightly reduced |
| Scanner hardware modification | 20 | Yes | No | No | No change | No change |

^a The actual reduction of acoustic noise can vary substantially depending on the specific equipment and implementation; these numbers are provided as a rough estimate.



may be acceptable. From an imaging perspective continuous imaging will generally provide the largest quantity of data, and no special considerations are necessary when analyzing the data. Continuous EPI scanning has been used in countless studies to identify brain networks responding to environmental sounds, speech, and music. The critical question is whether the cognitive processes being imaged are actually the ones in which the experimenter is interested².

SPARSE IMAGING

When researchers are concerned about acoustic noise in fMRI, by far the most widely used approach is sparse imaging, also referred to as clustered volume acquisition (Scheffler et al., 1998; Eden et al., 1999; Edmister et al., 1999; Hall et al., 1999; Talavage and Hall, 2012). In sparse imaging, illustrated in **Figure 4B**, the repetition time (TR) is set to be longer than the acquisition time (TA) of a single volume. Slice acquisition is clustered toward the end of a TR, leaving a period in which no data are collected. This intervening period is relatively quiet due to the lack of gradient switching, and permits stimuli to be presented in more favorable acoustic conditions. Because of the inherent lag of the hemodynamic response (typically 4–7 s to peak), the scan following stimulus presentation can still measure responses to stimuli, including the peak response if presentation is timed appropriately.

²For researchers who question whether acoustic noise during fMRI may impact cognitive processing, it may be interesting to suggest to a cognitive psychologist that they play 100 dB SPL sounds during their next behavioral study and gauge their enthusiasm.

The primary disadvantage of sparse imaging is that due to the longer TR, less information is available about the timecourse of the response (i.e., there is a lower sampling rate). In addition to reducing the accuracy of the response estimate, the reduced sampling rate also means that differences in timing of response may be interpreted as differences in magnitude. An example of this is shown in **Figure 4B**, in which hemodynamic responses that differ in magnitude and timing will give different results, depending on the time at which the response is sampled.

The lack of timecourse information in sparse imaging can be ameliorated in part by systematically varying the delay between the stimulus and volume collection (Robson et al., 1998; Belin et al., 1999), illustrated in **Figure 4D**. In this way, the hemodynamic response can be sampled at multiple time points relative to stimulus onset over different trials. Thus, across trials, an accurate temporal profile for each category of stimulus can be estimated. Like all event-related fMRI analyses this approach assumes a consistent response for all stimuli in a given category. It also may require prohibitively long periods of scanning to sample each stimulus at multiple points; this requirement has meant that in practice varying presentation times relative to data collection is done infrequently.

Many studies incorporating sparse imaging use an event-related design, along with TRs in the neighborhood of 16 s or greater, in order to allow scanner-induced BOLD response to return to near baseline levels on each trial. Although this may be particularly helpful for experiments in which activity in primary auditory areas is of interest, it is not necessary for all studies, and in principle sparse designs can use significantly shorter TRs (e.g., <5 s). Sometimes referred to as “fast” sparse designs, sparse designs with shorter TRs enable researchers to take advantage of a faster stimulus presentation rate and acquire more data for a given period of time, and for many experiments may be a more efficient approach (Perrachione and Ghosh, 2013).

Cardiac gating

Cardiac gating addresses problems caused by the fact that heart-beat and associated changes in blood flow can displace brainstem structures, making activity in these regions difficult to detect. With cardiac gating, researchers monitor a subject’s heart rate, and then adjust volume acquisition to be synchronized to the heart rate (i.e., occurring at a consistent time in the heart rate cycle) (Guimaraes et al., 1998). Because heart rate will not perfectly align with a chosen TR, using cardiac gating results in a variable TR (\pm approximately $\frac{1}{2}$ heart rate). (With relatively long TRs, the variability in sampling rate is typically not a significant problem, as the response to one trial is unlikely to overlap the response to another trial). Cardiac gating reduces data variability due to cardiac pulse motion artifacts and can thus improve ability to detect activity in subcortical structures prone to these artifacts, such as the inferior colliculus and medial geniculate body (Harms and Melcher, 2002; Overath et al., 2012).

INTERLEAVED SILENT STEADY STATE (ISSS) IMAGING

The main disadvantages in traditional sparse imaging come from the lack of information about the timecourse of the hemodynamic response, and the relatively small amount of data collected

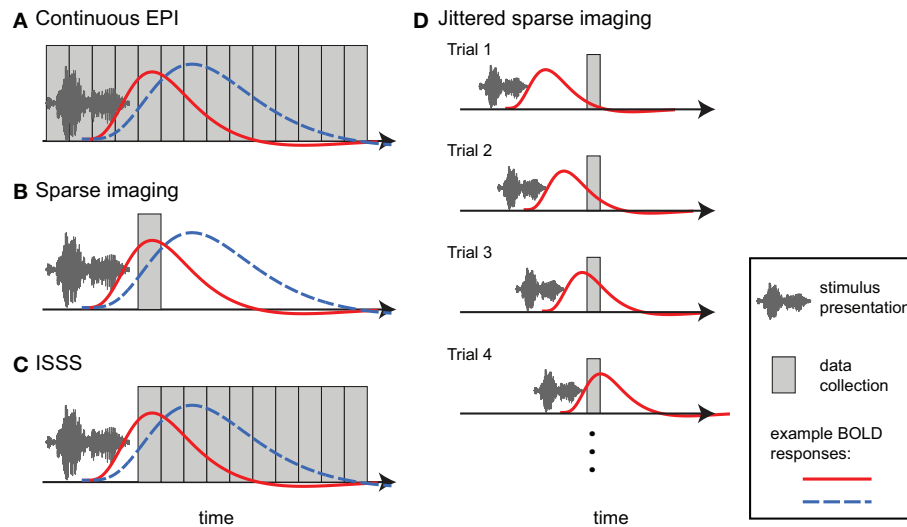


FIGURE 4 | Different approaches to imaging auditory stimuli provide varying compromises between temporal resolution and acoustic noise. Example BOLD responses are shown in blue and red; these could reflect different responses across individuals or experimental conditions. **(A)** Continuous EPI provides relatively good temporal resolution, but with a high level of continuous acoustic noise. **(B)** Sparse imaging includes a period in which no data is collected, allowing the presentation of stimuli in relative quiet (due to the absence of gradient switching noise). The delay in the hemodynamic response enables the peak response to be collected after stimulus presentation has finished. The reduced temporal resolution of a traditional sparse imaging sequence may obscure differences in response latency or

shape. In the hypothetical example, the blue response peaks higher than the red response; however, at the time when the sparse data point is collected, the red response is higher. **(C)** With interleaved silent steady state (ISSS) imaging, stimuli can also be presented in the absence of gradient switching noise, but a greater amount of data can be collected after presentation compared to sparse imaging. The delay in the hemodynamic response enables peak responses to be collected with relatively good temporal resolution. **(D)** By varying the time at which stimuli are presented relative to data collection across trials, non-continuous imaging can still provide information about the timecourse of the average response to a category of stimuli. Note how a different part of the BOLD response is sampled on each trial.

(leading to potentially less accurate parameter estimates and fewer degrees of freedom in first-level analyses). Although in principle multiple volumes can be acquired following each silent period, the equilibrium state of the brain tissue changes during these silent periods: The additional scans do not reflect steady-state longitudinal magnetization, and thus vary over time. The lack of steady-state longitudinal magnetization adds variance to the data that can be challenging to account for in timeseries statistical models.

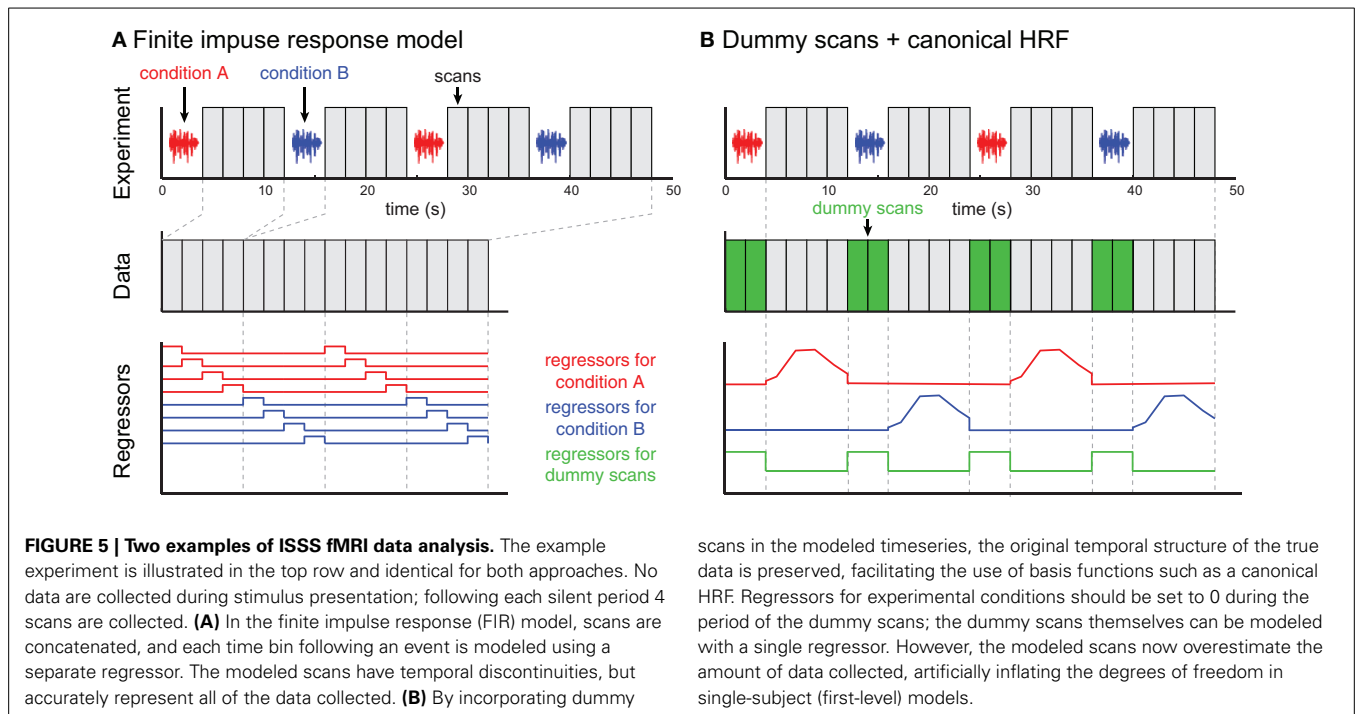
Schwarzbaumer et al. (2006) developed a solution to this problem by implementing a sequence with continuous excitation RF pulses, but variable readout gradients. The excitation pulses maintain steady state longitudinal magnetization but produce relatively little acoustic noise. As in traditional sparse imaging, an ISSS sequence permits stimuli to be presented in quiet and the peak BOLD activity to be captured due to the delay in hemodynamic response. However, with ISSS any number of volumes can be obtained following a silent period, as illustrated in **Figure 4C**. Although technically the temporal resolution is reduced relative to continuous scanning—as there are times when no data is being collected—the *effective* temporal resolution can be nearly as good as continuous scanning because data collection can capture much of the BOLD response following stimulus presentation: The ability of the sequence to capture the early hemodynamic response is limited solely by the length of the stimuli (with shorter stimuli permitting data collection to start closer to stimulus onset).

ISSS thus combines advantages of continuous and sparse imaging, allowing the presentation of stimuli in relative quiet, while still providing information on the timing of the hemodynamic response. Variations of ISSS fMRI have now been used successfully in numerous studies of auditory processing (Doehrmann et al., 2008, 2010; Bekinschtein et al., 2011; Davis et al., 2011; Engel and Keller, 2011; Mueller et al., 2011; Rodd et al., 2012; Yoo et al., 2012).

Compared to continuous or sparse imaging data, ISSS data can be challenging to analyze because the data are discontinuous—that is, the sampling rate is not consistent. Because of this added wrinkle, below I briefly review two examples of analyzing ISSS data, illustrated in **Figure 5**. No doubt with increasing experience ISSS data analysis can be further refined. These descriptions are based on an imaginary event-related fMRI study with two conditions (A and B) and a TR of 2 s. Each trial involves presenting a single stimulus during a period of 4 s of silence, followed by 8 s of data acquisition. With a TR of 2 s, this results in 4 volumes of data per trial.

Analyzing ISSS fMRI data using a finite impulse response (FIR) model

Perhaps the most straightforward approach to analyzing ISSS fMRI data is to use a finite impulse response (FIR) model, shown in **Figure 5A**. A typical FIR model would consider only the scans on which data was collected. The model would thus



have 4 regressors for condition A (one for each volume following stimulus presentation), and 4 regressors for condition B. These regressors would model the response at each time bin following a stimulus, making no assumptions about the shape of the response. As with any FIR Model, given the multiple regressors for each condition, there are several ways of summarizing the response to a condition, including an *F*-test over all 4 columns for a condition (asking: is there any effect at any time point?) or a *t*-test over all 4 columns (on average is there an increased response?). Similar options exist for comparing response between conditions.

Because the ISSS scans are not continuous, care must be taken when implementing temporal filtering, including typical highpass filtering done on fMRI timeseries data. Omitting highpass filtering may make an analysis particularly susceptible to the influence of low-frequency (non-acoustic) noise. One way to help mitigate this issue is to ensure trials of different conditions are not too far apart in time so that comparisons across conditions are not confounded with low-frequency fluctuations in the signal.

Analyzing ISSS fMRI data using dummy scans to mimic a continuous timeseries

An alternative approach is to ensure that rows of the design matrix correspond to a continuous timeseries, illustrated in **Figure 5B**. To accomplish this, dummy volumes can be included in the design matrix during the period in which no data were actually collected. A straightforward option is to use the mean EPI image across all (real) volumes in a session, although any identical image will work: Using an identical image for all dummy images means that all dummy images can be perfectly modeled using a single regressor (0 for real scans, 1 for dummy scans). With this model it is then possible to use a canonical HRF (or any other basis set) for

events of interest; the parameter estimates for these regressors are not influenced by the dummy scans. It is important to set the values for the non-dummy regressors to zero during the dummy scans to preserve estimation of the parameter estimate, and to rescale the regressors so that the maximum values are matched after these adjustments.

It is not actually necessary to use dummy scans in order to take advantage of timeseries properties, such as highpass filtering or using an informed basis function (e.g., a canonical HRF); an appropriate design matrix that takes into account the discontinuous nature of the data could be constructed. However, the use of dummy scans facilitates constructing design matrices within common fMRI analysis software packages, which are typically designed to work with continuous timeseries data.

When dummy scans are included in the final design matrix, the default degrees of freedom in the model will be incorrectly high, as the dummy scans should not be counted as observations. Thus, for first-level (single subject) analyses, an adjustment to the degrees of freedom should be made for valid statistical inference. For group analyses using a two-stage summary statistics procedure, however, adjusting for first-level degrees of freedom is not necessary.

ACTIVE NOISE CONTROL

A different approach to reducing the impact of acoustic noise in the MRI scanner is to change the way this sound is perceived by listeners using active noise control (Hall et al., 2009). As typically implemented, active noise control involves measuring the properties of the scanner noise, and generating a destructive acoustic signal (also known as “antinoise”) which is sent to the headphones that cancels a portion of the scanner noise (Chambers et al., 2001, 2007; Li et al., 2011). The destructive signal is based

on estimates of scanner noise that can either be fixed, or adjusted over the course of a scanning session to accommodate changes in the scanner noise. Adjusting over time may be important in the context of fMRI as subjects may move their heads over the course of a scanning session, which affects the acoustic characteristics of the noise reaching their ears.

In addition to sound presentation hardware, active noise control also requires an MR-compatible method for measuring the acoustic noise in the scanner, used to shape the destructive noise pulses. Whether sound is generated in the headset, or passed through a tube, the timing of this canceling sound is critical, as it must arrive with the specified phase relationship to the scanner noise.

Active noise control can reduce the level of acoustic noise by 30–40 dB, and subjective loudness by 20 dB (the difference between these measures likely reflecting the contribution of bone conducted vibration) (Hall et al., 2009; Li et al., 2011). Particularly relevant is that when using relatively simple auditory stimuli (pure tone pitch discrimination), (1) behavioral performance in the scanner was significantly better and (2) activity in primary auditory regions was significantly greater under conditions of active noise control compared to normal presentation (Hall et al., 2009).

USING CONTINUOUS fMRI SEQUENCES WITH REDUCED ACOUSTIC NOISE

Software modifications to EPI sequences intended to reduce the effects of acoustic scanner noise can be broadly grouped into two approaches: changing the nature of the acoustic stimulation and reducing the overall sound levels.

One approach to reducing sound levels of a standard EPI sequence is to modify the gradient pulse shape (Hennel et al., 1999). Typically, gradient pulses are trapezoidal, to increase the speed and efficiency of gradient encoding. By using sinusoidal pulses, acoustic noise can be reduced during BOLD fMRI (Loenneker et al., 2001), with some increase in the spatial smoothness of the reconstructed data.

Building on the idea of modified pulse shape, another type of quiet fMRI sequence was introduced by Schmitter et al. (2008). Their quiet continuous EPI sequence takes advantage of two key modifications to reduce acoustic noise. The first involves collecting data using a sinusoidal traversal of k space, enabling more gradual gradient switching (readout gradients are purely sinusoidal) and reducing the acoustic noise produced. The second modification addresses the fact that a large component of the acoustic noise during EPI comes from the resonance of the scanner hardware to the gradient switching. This reflects specific physical properties of each scanner, and varies across different speeds of gradient switching. Thus, it is possible to perform scanner-specific measurements of the acoustic noise generated for different readout gradient switching frequencies, and select a combination of parameters that is relatively quiet, but does not unacceptably compromise signal quality. In Peelle et al. (2010a), we chose a bandwidth of 1220 Hz/Px and an echo time of 44 ms (compared to a standard sequence with values of 2230 Hz/Px and 30 ms, respectively). As might be expected, the longer echo time lead to moderate increases in signal dropout in regions prone to

susceptibility artifact, such as inferior temporal and ventromedial frontal cortex. Together, these modifications produce approximately a 20 dB reduction in acoustic noise for the scanner, and using this sequence results in greater activity in several auditory regions compared to a standard continuous sequence (Schmitter et al., 2008; Peelle et al., 2010a).

Taking another approach to reducing the impact of scanner noise on observed activation, Seifritz et al. (2006) developed a continuous-sound EPI sequence to reduce the auditory stimulation caused by rapid acoustic pulses (Harms and Melcher, 2002), as found in conventional EPI. In their sequence the RF excitation pulses, phase-encoding gradients, and readout gradients are divided into short trains. The resulting repetition rate is fast enough that the acoustic noise is perceived as a continuous sound, rather than the pulsed sound perceived in conventional EPI. Using sparse imaging, the authors compared neural activity in response to audio recordings of conventional EPI compared to the “continuous sound” sequence. They found that the continuous sound sequence resulted in reduced activity in auditory cortex due to scanner noise, and increased activity to experimental manipulations.

SCANNER HARDWARE MODIFICATION

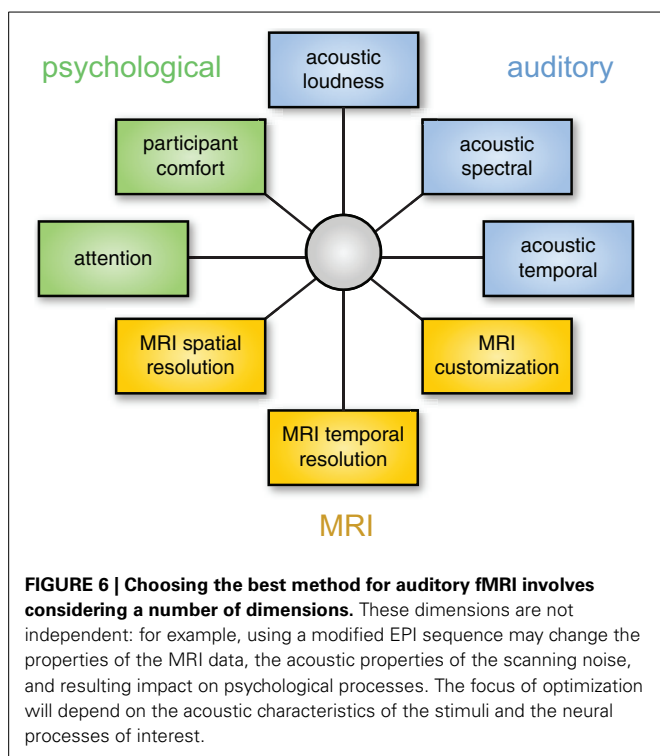
Although it may not be practical for most research groups to significantly modify scanner hardware, by changing the physical configuration of the MRI scanner it is possible to significantly reduce the amount of acoustic noise generated. Some approaches have included the use of rotating coils to reduce gradient switching (Cho et al., 1998b), placing the gradient coils in a vacuum to reduce noise propagation (Katsunuma et al., 2001), or altering the coil structure (Mansfield and Haywood, 2000). By combining multiple approaches and focusing on the largest contributors to acoustic noise, substantial reductions in noise levels can be achieved (Edelstein et al., 2002). In the future, commercial applications of these approaches may help to limit the impact of scanner noise during fMRI, particularly when combined with some of the other solutions outlined above.

AUDITORY fMRI IN NONHUMAN ANIMALS

Although my focus has been on fMRI of human subjects, many of these same challenges and solutions apply equally when using fMRI with animals (Petkov et al., 2009; Hall et al., 2014). As with human listeners, the choice of scanning protocol will depend on a researcher's primary interests and the acceptable level of tradeoff between data quality, temporal resolution, and acoustic noise. Although some concerns about attention and cognitive challenge may be mitigated when dealing with sedated animals, in the absence of empirical support it is probably not safe to assume that one protocol will prove optimal in all situations. In addition, the timing parameters of any non-continuous sequence will naturally need to be optimized for the HRF of the animal being studied (Brown et al., 2013).

CHOOSING THE APPROPRIATE SOLUTION

As discussed above, different solutions for auditory fMRI have intrinsic strengths and weaknesses, and thus any chosen approach involves a degree of compromise with respect to acoustic noise



(loudness or quality), psychological impact, and MRI data characteristics. It may be useful to think about this in a framework of multidimensional optimization, as illustrated in **Figure 6**. Because these dimensions are not independent, it is impossible to optimize for everything simultaneously (for example, approaches that have the lowest acoustic noise also tend to have poorer temporal resolution, forcing a researcher to choose between noise level and temporal resolution). It is therefore important to identify the dimensions that are most important for a given study. These will depend on the specific stimuli and scientific question at hand.

Although there are exceptions, as a general rule it is probably safest to prioritize the auditory and psychological aspects of data collection. If the processing of stimuli is affected by scanner noise (through masking or increased perceptual effort), the resulting neural processing may differ from what the researchers are interested in. In this case increased image quality will not help in identifying neural activity of interest. Thus, a sparse imaging sequence is nearly always preferable to continuous sequences because it presents the lowest level of background noise, and is straightforward to implement. If possible, an ISSS sequence presents an even stronger solution as it permits the presentation of stimuli in relative quiet, while not sacrificing temporal resolution to the same degree as a traditional sparse sequence.

When it is not feasible to present stimuli in the absence of scanner noise, considering the acoustic characteristics of the stimuli is critical. For example, if speech prosody, voice/speaker perception, or musical timbre is of interest, spectral cues may be particularly important, and thus the spectrum of the scanner noise may be a deciding factor. In contrast, for other stimuli (such as musical

beat, or other aspects of speech perception) temporal factors may dominate.

That being said, from a practical standpoint the majority of researchers will be constrained by available sequences and equipment, and thus the most common choice will be between a continuous EPI sequence and a traditional sparse sequence. In this case, adapting a paradigm and stimuli to work with the sparse sequence is almost always a safer choice.

RELYING ON CONVERGING EVIDENCE TO SUPPORT CONCLUSIONS

Although it is no doubt important to optimize fMRI acquisition and analysis parameters for auditory studies, the strongest inferences will always be drawn based on converging evidence from multiple modalities. With respect to auditory processing, this includes functional neuroimaging methods that allow the measuring of neural response in the absence of external noise such as positron emission tomography (PET), electroencephalography (EEG), magnetoencephalography (MEG), electrocorticography (ECoG), or optical imaging, as well as studying behavior in people with differences in brain structure (e.g., as a result of stroke or neurodegenerative disease).

CONCLUSIONS AND RECOMMENDATIONS

Echoplanar fMRI is acoustically noisy and poses considerable challenges for researchers interested in studying auditory processing. Although it is impossible to fully resolve the tension between the acoustic noise produced during fMRI and the desired experimental environment, the following steps will often be helpful in optimizing auditory responses and our interpretation of them:

- (1) Address, rather than ignore, the possible effects of background noise on activity seen in fMRI studies. Considering scanner noise is particularly important when using auditory stimuli, but may apply to non-auditory stimuli as well.
- (2) When possible, use methods that limit the impact of acoustic noise during fMRI scanning.
- (3) Provide empirical demonstrations of the effect of scanner noise on specific paradigms and analyses.

It is an exciting time for auditory neuroscience, and continuing technical and methodological advances suggest an even brighter (though hopefully quieter) future.

ACKNOWLEDGMENTS

Research reported in this publication was supported by the Dana Foundation and the National Institute on Aging of the National Institutes of Health under award number R01AG038490. I am grateful to Rhodri Cusack for helpful comments on this manuscript.

REFERENCES

- Adank, P. (2012). The neural bases of difficult speech comprehension and speech production: two Activation Likelihood Estimation (ALE) meta-analyses. *Brain Lang.* 122, 42–54. doi: 10.1016/j.bandl.2012.04.014
- Aguirre, G. K., Zarahn, E., and D'Esposito, M. (1998). The variability of human, BOLD hemodynamic responses. *Neuroimage* 8, 360–369. doi: 10.1006/nimg.1998.0369

- Amaro, E., Williams, S. C., Shergill, S. S., Fu, C. H. Y., MacSweeney, M., Picchioni, M. M., et al. (2002). Acoustic noise and functional magnetic resonance imaging: current strategies and future prospects. *J. Magn. Reson. Imaging* 16, 497–510. doi: 10.1002/jmri.10186
- Bandettini, P. A., Jesmanowicz, A., Van Kylen, J., Birn, R. M., and Hyde, J. S. (1998). Functional MRI of brain activation induced by scanner acoustic noise. *Magn. Reson. Med.* 39, 410–416. doi: 10.1002/mrm.1910390311
- Bekinschtein, T. A., Davis, M. H., Rodd, J. M., and Owen, A. M. (2011). Why clowns taste funny: the relationship between humor and semantic ambiguity. *J. Neurosci.* 31, 9665–9671. doi: 10.1523/JNEUROSCI.5058-10.2011
- Belin, P., Zatorre, R. J., Hoge, R., Evans, A. C., and Pike, B. (1999). Event-related fMRI of the auditory cortex. *Neuroimage* 10, 417–429. doi: 10.1006/nimg.1999.0480
- Besser, J., Zekveld, A. A., Kramer, S. E., Rönnerberg, J., and Festen, J. M. (2012). New measures of masked text recognition in relation to speech-in-noise perception and their associations with age and cognitive abilities. *J. Speech Lang. Hear. Res.* 55, 194–209. doi: 10.1044/1092-4388(2011/11-0008)
- Brown, T. A., Joanisse, M. F., Gati, J. S., Hughes, S. M., Nixon, P. L., Menon, R. S., et al. (2013). Characterization of the blood-oxygen level-dependent (BOLD) response in cat auditory cortex using high-field fMRI. *Neuroimage* 64, 458–465. doi: 10.1016/j.neuroimage.2012.09.034
- Chambers, J., Akeroyd, M. A., Summerfield, A. Q., and Palmer, A. R. (2001). Active control of the volume acquisition noise in functional magnetic resonance imaging: method and psychoacoustical evaluation. *J. Acoust. Soc. Am.* 110, 3041–3054. doi: 10.1121/1.1408948
- Chambers, J., Bullock, D., Kahana, Y., Kots, A., and Palmer, A. (2007). Developments in active noise control sound systems for magnetic resonance imaging. *Appl. Acoust.* 68, 281–295. doi: 10.1016/j.apacoust.2005.10.008
- Cho, Z.-H., Chung, S.-C., Lim, D.-W., and Wong, E. K. (1998a). Effects of the acoustic noise of the gradient systems on fMRI: a study on auditory, motor, and visual cortices. *Magn. Reson. Med.* 39, 331–335. doi: 10.1002/mrm.1910390224
- Cho, Z.-H., Chung, S. T., Chung, J. Y., Park, S. H., Kim, J. S., Moon, C. H., et al. (1998b). A new silent magnetic resonance imaging using a rotating DC gradient. *Magn. Reson. Med.* 39, 317–321. doi: 10.1002/mrm.1910390221
- Cousins, K. A. Q., Dar, H., Wingfield, A., and Miller, P. (2014). Acoustic masking disrupts time-dependent mechanisms of memory encoding in word-list recall. *Mem. Cognit.* 42, 622–638. doi: 10.3758/s13421-013-0377-7
- Cusack, R., Cumming, N., Bor, D., Norris, D., and Lyzenga, J. (2005). Automated post-hoc noise cancellation tool for audio recordings acquired in an MRI scanner. *Hum. Brain Mapp.* 24, 299–304. doi: 10.1002/hbm.20085
- Davis, M. H., Ford, M. A., Kherif, F., and Johnsrude, I. S. (2011). Does semantic context benefit speech understanding through “top-down” processes? Evidence from time-resolved sparse fMRI. *J. Cogn. Neurosci.* 23, 3914–3932. doi: 10.1162/jocn_a_00084
- Davis, M. H., and Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *J. Neurosci.* 23, 3423–3431.
- Davis, M. H., and Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hear. Res.* 229, 132–147. doi: 10.1016/j.heares.2007.01.014
- Doehrmann, O., Naumer, M. J., Volz, S., Kaiser, J., and Altmann, C. F. (2008). Probing category selectivity for environmental sounds in the human auditory brain. *Neuropsychologia* 46, 2776–2786. doi: 10.1016/j.neuropsychologia.2008.05.011
- Doehrmann, O., Weigelt, S., Altmann, C. F., Kaiser, J., and Naumer, M. J. (2010). Audiovisual functional magnetic resonance imaging adaptation reveals multi-sensory integration effects in object-related sensory cortices. *J. Neurosci.* 30, 3370–3379. doi: 10.1523/JNEUROSCI.5074-09.2010
- Dosenbach, N. U. F., Fair, D. A., Cohen, A. L., Schlaggar, B. L., and Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends Cogn. Sci.* 12, 99–105. doi: 10.1016/j.tics.2008.01.001
- Eckert, M. A., Menon, V., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., et al. (2009). At the heart of the ventral attention system: the right anterior insula. *Hum. Brain Mapp.* 30, 2530–2541. doi: 10.1002/hbm.20688
- Edelstein, W. A., Hede, R. A., Mallozzi, R. P., El-Hamamsy, S.-A., Ackermann, R. A., and Havens, T. J. (2002). Making MRI quieter. *Magn. Reson. Imaging* 20, 155–163. doi: 10.1016/S0730-725X(02)00475-7
- Eden, G. F., Joseph, J. E., Brown, H. E., Brown, C. P., and Zeffiro, T. A. (1999). Utilizing hemodynamic delay and dispersion to detect fMRI signal change without auditory interference: the behavior interleaved gradients technique. *Magn. Reson. Med.* 41, 13–20.
- Edmister, W. B., Talavage, T. M., Ledden, P. J., and Weisskoff, R. M. (1999). Improved auditory cortex imaging using clustered volume acquisitions. *Hum. Brain Mapp.* 7, 89–97.
- Elliott, M. R., Bowtell, R. W., and Morris, P. G. (1999). The effect of scanner sound in visual, motor, and auditory functional MRI. *Magn. Reson. Med.* 41, 1230–1235.
- Engel, A., and Keller, P. E. (2011). The perception of musical spontaneity in improvised and imitated jazz performances. *Front. Psychol.* 2:83. doi: 10.3389/fpsyg.2011.00083
- Erb, J., Henry, M. J., Eisner, F., and Obleser, J. (2013). The brain dynamics of rapid perceptual adaptation to adverse listening conditions. *J. Neurosci.* 33, 10688–10697. doi: 10.1523/JNEUROSCI.4596-12.2013
- Foster, J. R., Hall, D. A., Summerfield, A. Q., Palmer, A. R., and Bowtell, R. W. (2000). Sound-level measurements and calculations of safe noise dosage during EPI at 3 T. *J. Magn. Reson. Imaging* 12, 157–163. doi: 10.1002/1522-2586(200007)12:1<157::AID-JMRI17>3.0.CO;2-M
- Gaab, N., Gabrieli, J. D. E., and Glover, G. H. (2007). Assessing the influence of scanner background noise on auditory processing. II. An fMRI study comparing auditory processing in the absence and presence of recorded scanner noise using a sparse design. *Hum. Brain Mapp.* 28, 721–732. doi: 10.1002/hbm.20299
- Guimaraes, A. R., Melcher, J. R., Talavage, T. M., Baker, J. R., Ledden, P. J., Rosen, B. R., et al. (1998). Imaging subcortical auditory activity in humans. *Hum. Brain Mapp.* 6, 33–41.
- Hall, A. J., Brown, T. A., Grahn, J. A., Gati, J. S., Nixon, P. L., Hughes, S. M., et al. (2014). There's more than one way to scan a cat: imaging cat auditory cortex with high-field fMRI using continuous or sparse sampling. *J. Neurosci. Methods* 224, 96–106. doi: 10.1016/j.jneumeth.2013.12.012
- Hall, D. A., Chambers, J., Akeroyd, M. A., and Foster, J. R. (2009). Acoustic, psychophysical, and neuroimaging measurements of the effectiveness of active cancellation during auditory functional magnetic resonance imaging. *J. Acoust. Soc. Am.* 125, 347–359. doi: 10.1121/1.3021437
- Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., et al. (1999). “Sparse” temporal sampling in auditory fMRI. *Hum. Brain Mapp.* 7, 213–223.
- Hall, D. A., Summerfield, A. Q., Gonçalves, M. S., Foster, J. R., Palmer, A. R., and Bowtell, R. W. (2000). Time-course of the auditory BOLD response to scanner noise. *Magn. Reson. Med.* 43, 601–606. doi: 10.1002/(SICI)1522-2594(200004)43:4<601::AID-MRM16>3.0.CO;2-R
- Haller, S., Bartsch, A. J., Radue, E. W., Klarhöfer, M., Seifritz, E., and Scheffler, K. (2005). Effect of fMRI acoustic noise on non-auditory working memory task: comparison between continuous and pulsed sound emitting EPI. *MAGMA* 18, 263–271. doi: 10.1007/s10334-005-0010-2
- Haller, S., Hornola, G. A., Scheffler, K., Beckmann, C. F., and Bartsch, A. J. (2009). Background MR gradient noise and non-auditory BOLD activations: a data-driven perspective. *Brain Res.* 1282, 74–83. doi: 10.1016/j.brainres.2009.05.094
- Harms, M. P., and Melcher, J. R. (2002). Sound repetition rate in the human auditory pathway: representations in the waveshape and amplitude of fMRI activation. *J. Neurophysiol.* 88, 1433–1450. doi: 10.1152/jn.00156.2002
- Heinrich, A., and Schneider, B. A. (2011). Elucidating the effects of ageing on remembering perceptually distorted word pairs. *Q. J. Exp. Psychol.* 64, 186–205. doi: 10.1080/17470218.2010.492621
- Hennel, F., Girard, F., and Loenneker, T. (1999). “Silent” MRI with soft gradient pulses. *Magn. Reson. Med.* 42, 6–10.
- Hommel, B., Fischer, R., Colzato, L. S., van den Wildenberg, W. P. M., and Cellini, C. (2012). The effect of fMRI (noise) on cognitive control. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 290–301. doi: 10.1037/a0026353
- Junqua, J.-C. (1996). The influence of acoustics on speech production: a noise-induced stress phenomenon known as the Lombard reflex. *Speech Commun.* 20, 13–22. doi: 10.1016/S0167-6393(96)00041-6
- Katsunuma, A., Takamori, H., Sakakura, Y., Hamamura, Y., Ogo, Y., and Katayama, R. (2001). Quiet MRI with novel acoustic noise reduction. *MAGMA* 13, 139–144. doi: 10.1007/BF02678588
- Kuchinsky, S. E., Ahlstrom, J. B., Vaden, K. I. Jr., Cute, S. L., Humes, L. E., Dubno, J. R., et al. (2013). Pupil size varies with word listening and response selection difficulty in older adults with hearing loss. *Psychophysiology* 50, 23–34. doi: 10.1111/j.1469-8986.2012.01477.x

- Langers, D. R. M., and van Dijk, P. (2012). Mapping the tonotopic organization in human auditory cortex with minimally salient acoustic stimulation. *Cereb. Cortex* 22, 2024–2038. doi: 10.1093/cercor/bhr282
- Langers, D. R. M., van Dijk, P., and Backes, W. H. (2005). Interactions between hemodynamic responses to scanner acoustic noise and auditory stimuli in functional magnetic resonance imaging. *Magn. Reson. Med.* 53, 49–60. doi: 10.1002/mrm.20315
- Lash, A., Rogers, C. S., Zoller, A., and Wingfield, A. (2013). Expectation and entropy in spoken word recognition: effects of age and hearing acuity. *Exp. Aging Res.* 39, 235–253. doi: 10.1080/0361073X.2013.779175
- Li, M., Rudd, B., Lim, T. C., and Lee, J.-H. (2011). *In situ* active control of noise in a 4 T MRI scanner. *J. Magn. Reson. Imaging* 34, 662–669. doi: 10.1002/jmri.22694
- Loenneker, T., Hennel, F., Ludwig, U., and Hennig, J. (2001). Silent BOLD imaging. *MAGMA* 13, 76–81. doi: 10.1007/BF02668155
- Mansfield, P., and Haywood, B. (2000). Principles of active acoustic control in gradient coil design. *MAGMA* 10, 147–151. doi: 10.1007/BF02601849
- Mazard, A., Bazoyer, B., Etard, O., Tzourio-Mazoyer, N., Kosslyn, S. M., and Mellet, E. (2002). Impact of fMRI acoustic noise on the functional anatomy of visual mental imagery. *J. Cogn. Neurosci.* 14, 172–186. doi: 10.1162/08992902317236821
- McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R., and Wingfield, A. (2005). Hearing loss and perceptual effort: downstream effects on older adults' memory for speech. *Q. J. Exp. Psychol.* 58, 22–33. doi: 10.1080/02724980443000151
- Meléndez, J. C., and McCrank, E. (1993). Anxiety-related reactions associated with magnetic resonance imaging examinations. *JAMA* 270, 745–747. doi: 10.1001/jama.1993.03510060091039
- Miller, P., and Wingfield, A. (2010). Distinct effects of perceptual quality on auditory word recognition, memory formation and recall in a neural model of sequential memory. *Front. Syst. Neurosci.* 4:14. doi: 10.3389/fnsys.2010.00014
- Moelker, A., and Pattynama, P. M. T. (2003). Acoustic noise concerns in functional magnetic resonance imaging. *Hum. Brain Mapp.* 20, 123–141. doi: 10.1002/hbm.10134
- Moelker, A., Wielopolski, P. A., and Pattynama, P. M. T. (2003). Relationship between magnetic field strength and magnetic-resonance-related acoustic noise levels. *MAGMA* 16, 52–55. doi: 10.1007/s10334-003-0005-9
- Mueller, K., Mildner, T., Fritz, T., Lepsien, J., Schwarzbauer, C., Schroeter, M. L., et al. (2011). Investigating brain response to music: a comparison of different fMRI acquisition schemes. *Neuroimage* 54, 337–343. doi: 10.1016/j.neuroimage.2010.08.029
- Murphy, D. R., Craik, F. I. M., Li, K. Z. H., and Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychol. Aging* 15, 323–334. doi: 10.1037/0882-7974.15.2.323
- Obleser, J., and Kotz, S. A. (2010). Expectancy constraints in degraded speech modulate the language comprehension network. *Cereb. Cortex* 20, 633–640. doi: 10.1093/cercor/bhp128
- Obleser, J., Wise, R. J. S., Dresner, M. A., and Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *J. Neurosci.* 27, 2283–2289. doi: 10.1523/JNEUROSCI.4663-06.2007
- Overath, T., Zhang, Y., Sanes, D. H., and Poeppel, D. (2012). Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: fMRI evidence. *J. Neurophysiol.* 107, 2042–2056. doi: 10.1152/jn.00308.2011
- Peelle, J. E., and Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Front. Psychol.* 3:320. doi: 10.3389/fpsyg.2012.00320
- Peelle, J. E., Eason, R. J., Schmitter, S., Schwarzbauer, C., and Davis, M. H. (2010a). Evaluating an acoustically quiet EPI sequence for use in fMRI studies of speech and auditory processing. *Neuroimage* 52, 1410–1419. doi: 10.1016/j.neuroimage.2010.05.015
- Peelle, J. E., Johnsrude, I. S., and Davis, M. H. (2010b). Hierarchical processing for speech in human auditory cortex and beyond. *Front. Hum. Neurosci.* 4:51. doi: 10.3389/fnhum.2010.00051
- Perrachione, T. K., and Ghosh, S. S. (2013). Optimized design and analysis of sparse-sampling fMRI experiments. *Front. Neurosci.* 7:55. doi: 10.3389/fnins.2013.00055
- Petkov, C. I., Kayser, C., Augath, M., and Logothetis, N. K. (2009). Optimizing the imaging of the monkey auditory cortex: sparse vs. continuous fMRI. *Magn. Reson. Imaging* 27, 1065–1073. doi: 10.1016/j.mri.2009.01.018
- Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282
- Price, D. L., De Wilde, J. P., Papadaki, A. M., Curran, J. S., and Kitney, R. I. (2001). Investigation of acoustic noise on 15 MRI scanners from 0.2 T to 3 T. *J. Magn. Reson. Imaging* 13, 288–293. doi: 10.1002/1522-2586(200102)13:2<288::AID-JMRI1041>3.0.CO;2-P
- Pripfl, J., Robinson, S., Leodolter, U., Moser, E., and Bauer, H. (2006). EEG reveals the effect of fMRI scanner noise on noise-sensitive subjects. *Neuroimage* 31, 332–341. doi: 10.1016/j.neuroimage.2005.11.031
- Quirk, M., Letendre, A., Cliottone, R., and Lingley, J. (1989). Anxiety in patients undergoing MR imaging. *Radiology* 170, 463–466. doi: 10.1148/radiology.170.2.2911670
- Rabbitt, P. M. A. (1968). Channel capacity, intelligibility and immediate memory. *Q. J. Exp. Psychol.* 20, 241–248. doi: 10.1080/14640746808400158
- Ravicz, M. E., and Melcher, J. R. (2001). Isolating the auditory system from acoustic noise during functional magnetic resonance imaging: examination of noise conduction through the ear canal, head, and body. *J. Acoust. Soc. Am.* 109, 216–231. doi: 10.1121/1.1326083
- Ravicz, M. E., Melcher, J. R., and Kiang, N. Y.-S. (2000). Acoustic noise during functional magnetic resonance imaging. *J. Acoust. Soc. Am.* 108, 1683–1696. doi: 10.1121/1.1310190
- Robson, M. D., Dorosz, J. L., and Gore, J. C. (1998). Measurements of the temporal fMRI response of the human auditory cortex to trains of tones. *Neuroimage* 7, 185–198. doi: 10.1006/nimg.1998.0322
- Rodd, J. M., Johnsrude, I. S., and Davis, M. H. (2012). Dissociating frontotemporal contributions to semantic ambiguity resolution in spoken sentences. *Cereb. Cortex* 22, 1761–1773. doi: 10.1093/cercor/bhr252
- Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The ease of language understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031
- Scheffler, K., Bilecen, D., Schmid, N., Tschopp, K., and Seelig, J. (1998). Auditory cortical responses in hearing subjects and unilateral deaf patients as detected by functional magnetic resonance imaging. *Cereb. Cortex* 8, 156–163. doi: 10.1093/cercor/8.2.156
- Schmitter, S., Diesch, E., Amann, M., Kroll, A., Moayer, M., and Schad, L. R. (2008). Silent echo-planar imaging for auditory fMRI. *MAGMA* 21, 317–325. doi: 10.1007/s10334-008-0132-4
- Schwarzbauer, C., Davis, M. H., Rodd, J. M., and Johnsrude, I. (2006). Interleaved silent steady state (ISSS) imaging: a new sparse imaging method applied to auditory fMRI. *Neuroimage* 29, 774–782. doi: 10.1016/j.neuroimage.2005.08.025
- Scott, S. K., and McGettigan, C. (2013). The neural processing of masked speech. *Hear. Res.* 303, 58–66. doi: 10.1016/j.heares.2013.05.001
- Seifritz, E., Di Salle, F., Esposito, F., Herdener, M., Neuhoof, J. G., and Scheffler, K. (2006). Enhancing BOLD response in the auditory system by neurophysiologically tuned fMRI sequence. *Neuroimage* 29, 1013–1022. doi: 10.1016/j.neuroimage.2005.08.029
- Shah, N. J., Jäncke, L., Grosse-Ruyken, M.-L., and Müller-Gärtner, H.-W. (1999). Influence of acoustic masking noise in fMRI of the auditory cortex during phonetic discrimination. *J. Magn. Reson. Imaging* 9, 19–25.
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science* 270, 303–304. doi: 10.1126/science.270.5234.303
- Skouras, S., Gray, M., Critchley, H., and Koelsch, S. (2013). fMRI scanner noise interaction with affective neural processes. *PLoS ONE* 8:e80564. doi: 10.1371/journal.pone.0080564
- Sohoglu, E., Peelle, J. E., Carlyon, R. P., and Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *J. Neurosci.* 32, 8443–8453. doi: 10.1523/JNEUROSCI.5069-11.2012
- Surprenant, A. M. (1999). The effect of noise on memory for spoken syllables. *Int. J. Psychol.* 34, 328–333. doi: 10.1080/002075999399648
- Talavage, T. M., and Edmister, W. B. (2004). Nonlinearity of fMRI responses in human auditory cortex. *Hum. Brain Mapp.* 22, 216–228. doi: 10.1002/hbm.20029
- Talavage, T. M., Edmister, W. B., Ledden, P. J., and Weisskoff, R. M. (1999). Quantitative assessment of auditory cortex responses induced by imager acoustic noise. *Hum. Brain Mapp.* 7, 79–88.

- Talavage, T. M., Gonzalez-Castillo, J., and Scott, S. K. (2014). Auditory neuroimaging with fMRI and PET. *Hear. Res.* 307, 4–15. doi: 10.1016/j.heares.2013.09.009
- Talavage, T. M., and Hall, D. A. (2012). How challenges in auditory fMRI led to general advancements for the field. *Neuroimage* 62, 641–647. doi: 10.1016/j.neuroimage.2012.01.006
- Tomasi, D., Caparelli, E. C., Chang, L., and Ernst, T. (2005). fMRI-acoustic noise alters brain activation during working memory tasks. *Neuroimage* 27, 377–386. doi: 10.1016/j.neuroimage.2005.04.010
- Tun, P. A., Benichov, J., and Wingfield, A. (2010). Response latencies in auditory sentence comprehension: effects of linguistic versus perceptual challenge. *Psychol. Aging* 25, 730–735. doi: 10.1037/a0019300
- Vaden, K. I. J., Kuchinsky, S. E., Cute, S. L., Ahlstrom, J. B., Dubno, J. R., and Eckert, M. A. (2013). The cingulo-opercular network provides word-recognition benefit. *J. Neurosci.* 33, 18979–18986. doi: 10.1523/JNEUROSCI.1417-13.2013
- Van Engen, K. J., Chandrasekaran, B., and Smiljanic, R. (2012). Effects of speech clarity on recognition memory for spoken sentences. *PLoS ONE* 7:e43753. doi: 10.1371/journal.pone.0043753
- Van Gerven, P. W. M., Vos, H., Van Boxtel, M. P. J., Janssen, S. A., and Miedema, H. M. E. (2009). Annoyance from environmental noise across the lifespan. *J. Acoust. Soc. Am.* 126, 187–194. doi: 10.1121/1.3147510
- Wild, C. J., Yusuf, A., Wilson, D., Peelle, J. E., Davis, M. H., and Johnsrude, I. S. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *J. Neurosci.* 32, 14010–14021. doi: 10.1523/JNEUROSCI.1528-12.2012
- Wingfield, A., Tun, P. A., and McCoy, S. L. (2005). Hearing loss in older adulthood: what it is and how it interacts with cognitive performance. *Curr. Dir. Psychol. Sci.* 14, 144–148. doi: 10.1111/j.0963-7214.2005.00356.x
- Yoo, S., Chung, J.-Y., Jeon, H.-A., Lee, K.-M., Kim, Y.-B., and Cho, Z.-H. (2012). Dual routes for verbal repetition: articulation-based and acoustic-phonetic codes for pseudoword and word repetition, respectively. *Brain Lang.* 122, 1–10. doi: 10.1016/j.bandl.2012.04.011
- Zekveld, A. A., George, E. L. J., Kramer, S. E., Goverts, S. T., and Houtgast, T. (2007). The development of the text reception threshold test: a visual analogue of the speech reception threshold test. *J. Speech Lang. Hear. Res.* 50, 576–584. doi: 10.1044/1092-4388(2007/040)
- Zekveld, A. A., and Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: insights from pupillometry. *Psychophysiology* 51, 277–284. doi: 10.1111/psyp.12151

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 25 April 2014; accepted: 29 July 2014; published online: 21 August 2014.

Citation: Peelle JE (2014) Methodological challenges and solutions in auditory functional magnetic resonance imaging. *Front. Neurosci.* 8:253. doi: 10.3389/fnins.2014.00253

This article was submitted to Brain Imaging Methods, a section of the journal *Frontiers in Neuroscience*.

Copyright © 2014 Peelle. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.